On Detecting Harmonic Oscillations

Anatoli Juditsky* Arkadi Nemirovski†

January 24, 2013

Abstract

In this paper, we focus on the following testing problem: assume that we are given observations of a real-valued signal along the grid $0, 1, ..., N-1$, corrupted by the standard Gaussian noise. We want to distinguish between two hypotheses: (a) the signal is a *nuisance* – a linear combination of d_n harmonic oscillations of given frequencies, and (b): signal is the sum of a nuisance and a linear combination of a given number d_s of harmonic oscillations with *unknown* frequencies, and such that the distance (measured in the uniform norm on the grid) between the signal and the set of nuisances is at least $\rho > 0$. We propose a computationally efficient test for distinguishing between (a) and (b) and show that its "resolution" (the smallest value of ρ for which (a) and (b) are distinguished with a given confidence $1-\alpha$) is $O\left(\sqrt{\ln(N/\alpha)/N}\right)$, with the hidden factor depending solely on d_n and d_s and independent of the frequencies in question. We show that this resolution, up to a factor which is polynomial in d_n, d_s and *logarithmic in* N, is the best possible under circumstances. We further extend the outlined results to the case of nuisances and signals *close* to linear combinations of harmonic oscillations, and provide illustrative numerical results.

1 Introduction

In this paper, we address the detection problem as follows. A signal (two-sided sequence of reals) x is observed on time horizon $0, 1, ..., N-1$ according to

$$y = x_0^{N-1} + \xi,$$

where $\xi \sim \mathcal{N}(0, I_N)$ is the white Gaussian noise and $z_0^{N-1} = [z_0; ...; z_{N-1}]$. Given y we want to distinguish between two hypotheses:

- *Nuisance hypothesis:* $x \in H_0$, where H_0 is comprised of all *nuisances* – linear combinations of d_n harmonic oscillations of given frequencies;

*LJK, Université J. Fourier, B.P. 53, 38041 Grenoble Cedex 9, France, `Anatoli.Juditsky@imag.fr`

†Georgia Institute of Technology, Atlanta, Georgia 30332, USA, `nemirovs@isye.gatech.edu`

Research of the second author was supported by the ONR grant N000140811104.

- *Signal hypothesis:* $x \in H_1(\rho)$, where $H_1(\rho)$ is the set of all sequences x representable as $s+u$ with the "nuisance component" u belonging to H_0 and the "signal component" s being a sum of at most d_s harmonic oscillations (of whatever frequencies) such that the *uniform* distance, on the time horizon in question, from x to all nuisance signals is at least ρ:

$$\min_{z\in H_0} \|x_0^{N-1} - z_0^{N-1}\|_\infty \geq \rho.$$

We are interested in a test which allows to distinguish, with a given confidence $1-\alpha$, between the above two hypotheses for as small "resolution" ρ as possible.

An approach to this problem which is generally advocated in the signal processing literature is based on frequency estimation. The spectrum of the signal is first estimated using noise subspace methods, such as multiple signal classification (MUSIC) [10, 11], then the nuisance spectrum is removed from this estimation and the decision is taken whether the remaining "spectral content" indicates the presence of a signal or is a noise artifact (for detailed presentation of these techniques see [12, 5]). To the best of our knowledge, no theoretical bounds for the resolution of such tests are available. A different test for the case when no nuisance is present, based on the normalized periodogram, has been proposed in [3]. The properties of this test and of its various modifications were extensively studied in the statistical literature (see, e.g., [13, 4, 1, 5]). However, theoretical results on the power of this test are limited exclusively to the case of sequence x being a linear combination of Fourier harmonics $e^{2\pi\imath kt/N}$, $k = 0, 1, ..., N-1$ under signal hypothesis.

In this paper we show that a good solution to the outlined problem is offered by an extremely simple test as follows.

> *Let* $F_N u = \left\{\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1} u_t \exp\{2\pi\imath kt/N\}\right\}_{k=0}^{N-1} : \mathbf{C}^N \to \mathbf{C}^N$ *be the Discrete Fourier Transform. Given the observation y, we solve the convex optimization problem*
> $$\mathrm{Opt}(y) = \min_z \left\{\|F_N(y - z_0^{N-1})\|_\infty : z \in H_0\right\}$$
> *and compare the optimal value with a threshold $q_N(\alpha)$ which is a valid upper bound on the $1-\alpha$-quantile of $\|F_N\xi\|_\infty$, $\alpha \in (0,1)$ being a given tolerance:*
> $$\mathrm{Prob}_{\xi\sim\mathcal{N}(0,I_N)}\left\{\|F_N\xi\|_\infty > q_N(\alpha)\right\} \leq \alpha.$$
> *If $\mathrm{Opt}(y) \leq q_N(\alpha)$, we accept the nuisance hypothesis, otherwise we claim that a signal is present.*

It is immediately seen that the outlined test rejects the nuisance hypothesis when it is true with probability at most α [1]. Our main result (Theorems 3.1) states that *the probability to reject the signal hypothesis when it is true is $\leq \alpha$, provided that the resolution ρ is not too small, specifically,*

$$\rho \geq C(d_n + d_s)\sqrt{\ln(N/\alpha)/N} \tag{!}$$

[1] This fact is completely independent of what the nuisance hypothesis is – it remains true when H_0 is an arbitrary set in the space of signals.

with an appropriately chosen universal function $C(\cdot)$.

Some comments are in order.

- We show that in our detection problem the power of our test is nearly as good as it can be: precisely, for every pair d_n, d_s and properly selected H_0, no test can distinguish $(1-\alpha)$-reliably between H_0 and $H_1(\rho)$ when $\rho < O(1)d_s\sqrt{\ln(1/\alpha)/N}$ from now on, $O(1)$'s are appropriately chosen positive absolute constants.

- We are measuring the resolution in the "weakest" of all natural scales, namely, via the *uniform* distance from the signal to the set of nuisances. When passing from the uniform norm to the normalized Euclidean norm $|x_0^{N-1}|_2 = \|x_0^{N-1}\|_2/\sqrt{N} \leq \|x_0^N\|_\infty$, an immediate lower bound on the resolution which allows for reliable detection becomes $O(1)\sqrt{\ln(1/\alpha)/N}$. In the case when, as in our setting, signals obeying H_0 and $H_1(\rho)$ admit parametric description involving K parameters, this lower bound, up to a factor logarithmic in N and linear in K, is also an upper resolution bound, and the associated test is based on estimating the Euclidean distance from the signal underlying the observations to the nuisance set. Note that, in general, the $|\cdot|_2$-norm can be smaller than $\|\cdot\|_\infty$ by a factor as large as $\sqrt{N}$, and the fact that "energy-based" detection allows to distinguish well between parametric hypotheses "separated" by $O(\sqrt{\ln(N/\alpha)/N})$ in $|\cdot|_2$ norm does *not* automatically imply the possibility to distinguish between hypotheses separated by $O(\sqrt{\ln(N/\alpha)/N})$ in the uniform norm[2]. The latter possibility exists in the situation we are interested in due to the particular structure of our specific nuisance and non-nuisance hypotheses; this structure allows also for a dedicated *non-energy-based* test.

- For the sake of definiteness, throughout the paper we assume that the observation noise is the standard white Gaussian one. This assumption is by no means critical: *whatever is the observation noise, with $q_N(\alpha)$ defined as (an upper bound on) the $(1-\alpha)$-quantile of $\|F_N\xi\|_\infty$, the above test $(1-\alpha)$-reliably distinguishes between the hypotheses H_0 and $H_1(\rho)$, provided that $\rho \geq C(d_n+d_s)q_N(\alpha)/\sqrt{N}$.* For example, the results of Theorems 3.1 and 3.2 remain valid when the observation noise is of the form $\xi = \{\xi_t = \sum_{\tau=-\infty}^{\infty}\gamma_\tau\eta_{t-\tau}\}_{t=0}^{N-1}$ with deterministic γ_τ, $\sum_\tau|\gamma_\tau| \leq 1$, and independent $\eta_t \sim \mathcal{N}(0,1)$.

- The main observation underlying the results on the resolution of the above test is as follows: *when x is the sum of at most d harmonic oscillations, $\|F_N x\|_\infty \geq C(d)\sqrt{N}\|x_0^{N-1}\|_\infty$ with some universal positive function $C(d)$.* This observation originates from [8] and, along with its modifications and extensions, was utilized, for the time being in the denoising setting, in [9, 2, 6, 7]. It is worth to mention that it also allows to extend, albeit with degraded constants, the results of Theorems 3.1 and 3.2 to multi-dimensional setting.

[2] Indeed, let H_0 state that the signal is 0, and $H_1(\rho)$ state that the signal is $\geq \rho$ at $t=0$ and is zero for all other t's. These two hypotheses cannot be reliably distinguished unless $\rho \geq O(1)$, that is, the $\|\cdot\|_\infty$ resolution in this case is much larger than $O(\sqrt{\ln(N/\alpha)/N})$.

The rest of this paper is organized as follows. In section 2 we give a detailed description of the detection problems (P_1), (P_2), (N_1) and (N_2), we are interested in (where (P_2) is the problem we have discussed so far). Our test is presented in section 3 where we also provide associated resolution bounds for these problems. Next in section 4 we present lower bounds on "good" (allowing for $(1-\alpha)$-reliable hypotheses testing) resolutions, while in section 5 we present some numerical illustrations. The proofs of results of sections 3 and 4 are put into section 6.

2 Problem description

Let $\mathcal{S}$ stand for the space of all two-sided real sequences $z = \{z_t \in \mathbf{R}\}_{t=-\infty}^{\infty}$. Assume that a discrete time signal $x \in \mathcal{S}$ is observed on time horizon $0 \leq t < N$ according to

$$y = x_0^{N-1} + \xi,\ \xi \sim \mathcal{N}(0, I_N), \tag{1}$$

where (and from now on) for $z \in \mathcal{S}$ and integers $p \leq q$, z_p^q stands for the vector $[z_p; z_{p+1}; ...; z_q]$.

In the sequel, we are interested in the case when the signal is a linear combination of a given number of harmonic oscillations. Specifically, let Δ stand for the shift operator on $\mathcal{S}$:

$$(\Delta z)_t = z_{t-1}, \quad x \in \mathcal{S}.$$

Let Ω_d be the set of all unordered collections $\mathbf{w} = \{\omega_1, ..., \omega_d\}$ of d reals which are "symmetric mod 2π," meaning that for every a, the number of $\omega_i's$ equal, modulus 2π, to a is exactly the same as the number of ω_i's equal, modulus 2π, to $-a$. We associate with $\mathbf{w} \in \Omega_d$ the real algebraic polynomial

$$p_{\mathbf{w}}(\zeta) = \prod_{\ell=1}^{d} (1 - \exp\{\imath\omega_\ell\}\zeta)$$

and the subspace $\mathcal{S}[\mathbf{w}]$ of $\mathcal{S}$, comprised of $x \in \mathcal{S}$ satisfying the homogeneous finite-difference equation

$$p_{\mathbf{w}}(\Delta)x \equiv 0. \tag{2}$$

In other words, $\mathcal{S}[\{\omega_1, ..., \omega_d\}]$ is comprised of all real two-sided sequences of the form

$$x_t = \sum_{\ell=1}^{d} [p_\ell(t)\cos(\omega_\ell t) + q_\ell(t)\sin(\omega_\ell t)]$$

with real algebraic polynomials $p_\ell(\cdot)$, $q_\ell(\cdot)$ of degree $< m_\ell$, where m_ℓ is the multiplicity, $\bmod 2\pi$, of ω_ℓ in $\mathbf{w}$. We set

$$\mathcal{S}_d = \bigcup_{\mathbf{w} \in \Omega_d} \mathcal{S}[\mathbf{w}].$$

Remark 2.1 In what follows, we refer to the reals ω_i constituting $\mathbf{w} \in \Omega_d$ as to *frequencies* of a signal from $\mathcal{S}[\mathbf{w}]$. A reader would keep in mind that the number of "actual frequencies" in such a signal can be less than d: frequencies in $\mathbf{w}$ different from $0 \bmod 2\pi$ and $\pi \bmod 2\pi$ go in "symmetric pairs" $(\omega, \omega' = -\omega \bmod 2\pi)$, and such a pair gives rise to a single "actual frequency."

Given a positive integer N, real $\epsilon \geq 0$, and $\mathbf{w} \in \Omega_d$, we set

$$\mathcal{S}^{N,\epsilon}[\mathbf{w}] = \{x \in \mathcal{S} : \|[p_{\mathbf{w}}(\Delta)x]_0^{N-1}\|_\infty \leq \epsilon\}.$$

Finally, we denote $\mathcal{S}_d^{N,\epsilon}$ the set

$$\mathcal{S}_d^{N,\epsilon} = \bigcup_{\mathbf{w} \in \Omega_d} \mathcal{S}^{N,\epsilon}[\mathbf{w}].$$

When N is clear from the context, we shorten the notations $\mathcal{S}^{N,\epsilon}[\mathbf{w}]$, $\mathcal{S}_d^{N,\epsilon}$ to $\mathcal{S}^{\epsilon}[\mathbf{w}]$ and $\mathcal{S}_d^{\epsilon}$, respectively.

In the definitions above, it was tacitly assumed that d is a positive integer. It makes sense to allow also for the case of $d = 0$. By definition, Ω_0 is comprised of the empty collection $\mathbf{w} = \emptyset$ and $p_\emptyset(\zeta) \equiv 1$. With this convention, $\mathcal{S}^{N,\epsilon}[\emptyset] = \{x \in \mathcal{S} : \|x_0^{N-1}\| \leq \epsilon\}$.

Observe that the family $\mathcal{S}_d^{N,\epsilon}$, $d \geq 2$, is quite rich. For instance, it contains "smoothly varying signals" (case of $w_i = 0 \bmod 2\pi$), along with "fast varying" – amplitude-modulated and frequency-modulated signals (see [2, 6] for more examples).

We detail now the hypothesis testing problems about the sequence x via observation y given by (1). In what follows d_s and d_n are given positive integers, and ρ, ϵ_n, ϵ_s are given positive reals.

(P_1) The "basic" hypothesis testing problem we consider is that of testing of a simple nuisance hypothesis $\{x = 0\}$ against the alternative that a signal $x \in \mathcal{S}_{d_s}$ "is present," meaning that the uniform norm of the signal on the observation window $[0, ..., N-1]$ exceeds certain threshold $\rho > 0$. In other words, we consider the following set of hypotheses:

$$\begin{array}{rcl} H_0 &=& \{x = 0\}, \\ H_1(\rho) &=& \left\{x \in \mathcal{S}_{d_s} :\ \|x_0^{N-1}\|_\infty \geq \rho\right\}. \end{array}$$

(P_2) We suppose that $x \in \mathcal{S}$ decomposes into "signal" and "nuisance":

$$x = s + u,$$

where s is the signal of interest and a *nuisance* u belongs to a subspace $\mathcal{S}[\overline{\mathbf{w}}]$, assumed to be known *a priori*. We consider a composite nuisance hypothesis that x is a "pure nuisance", and the alternative (signal hypothesis) that useful signal s does not vanish, and the deviation, when measured in the uniform norm on the observation window, of "signal+nuisance" from the nuisance subspace is at least $\rho > 0$. Thus we arrive at the testing problem: given $\overline{\mathbf{w}} \in \Omega_{d_n}$ decide between the hypotheses

$$\begin{array}{rcl} H_0 &=& \{x = u \in \mathcal{S}[\overline{\mathbf{w}}]\}, \\ H_1(\rho) &=& \left\{\begin{array}{l} x = u + s :\ u \in \mathcal{S}[\overline{\mathbf{w}}],\ s \in \mathcal{S}_{d_s}, \\ \text{such that } \min_z\{\|[x - z]_0^{N-1}\|_\infty : z \in \mathcal{S}[\overline{\mathbf{w}}]\} \geq \rho. \end{array}\right\} \end{array}$$

Clearly, problem (P_1) is a particular case of (P_2) with $d_n = 0$ (and thus $\mathcal{S}_{d_n} = \{0\}$ is a singleton).

(N_1) Given $\overline{\mathbf{w}} \in \Omega_{d_n}$, consider the nonparametric nuisance hypothesis that the nuisance $u \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}]$ with some known $\overline{\mathbf{w}}$. The signal hypothesis is that the useful signal $s \in \mathcal{S}_{d_s}$ is present, and $x = s + u$ deviates from the nuisance set on the observation window by at least $\rho > 0$ in the uniform norm:

$$\begin{array}{rcl} H_0 & = & \{x = u \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}]\}, \\ H_1(\rho) & = & \left\{ \begin{array}{l} x = u + s : \ u \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}], \ s \in \mathcal{S}_{d_s}, \\ \text{such that } \min_z\{\|[x-z]_0^{N-1}\|_\infty : z \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}]\} \geq \rho. \end{array} \right\} \end{array}$$

(N_2) The last decision problem is a natural extension of (N_1): we consider the problem of testing a nonparametric nuisance hypothesis against a nonparametric signal alternative that the useful signal $s \in \mathcal{S}_{d_s}^{N,\epsilon_s}$ is present:

$$\begin{array}{rcl} H_0 & = & \{x = u \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}]\}, \\ H_1(\rho) & = & \left\{ \begin{array}{l} x = u + s : \ u \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}], \ s \in \mathcal{S}_{d_s}^{N,\epsilon_s}, \\ \text{and such that } \min_z\{\|[x-z]_0^{N-1}\|_\infty : z \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}]\} \geq \rho. \end{array} \right\} \end{array}$$

Note that problem (N_1) is a particular case of (N_2) with $\epsilon_s = 0$.

In the sequel, we refer to the sequences obeying H_0 (resp., $H_1 = H_1(\rho)$) as *nuisance* (resp. *signal*) sequences.

Let $\varphi(\cdot)$ be a test, i.e. a Borel function on $\mathbf{R}^N$ taking values in $\{0, 1\}$, which receives on input observation (1) (along with the data participating in the description of H_0 and H_1). The event $\{\varphi(y) = 1\}$ corresponds to rejecting the hypothesis H_0, while $\{\varphi(y) = 0\}$ implies that H_1 is rejected. The quality of the test is characterized by the error probabilities – the probabilities of rejecting erroneously each of the hypotheses:

$$\varepsilon_0(\varphi; H_0) = \sup_{x \in H_0} \mathrm{Prob}_x\{\varphi(y) = 1\}, \quad \varepsilon_1(\varphi; H_1(\rho)) = \sup_{x \in H_1(\rho)} \mathrm{Prob}_x\{\varphi(y) = 0\}.$$

We define the *risk of the test* as

$$\mathrm{Risk}(\varphi, \rho) = \max\left\{\varepsilon_0(\varphi; H_0), \varepsilon_1(\varphi; H_1(\rho))\right\}.$$

Let $\alpha \in (0, 1/2)$ be given. In this paper we address the following question: *for the testing problems above, what is the smallest possible ρ such that one can distinguish $(1-\alpha)$-reliably between the hypotheses H_0 and $H_1 = H_1(\rho)$ via observation* (1) (i.e., is such that $\mathrm{Risk}(\varphi, \rho) \leq \alpha$). In the sequel, we refer to such ρ as to *α-resolution* in the testing problem in question, and our goal is to find reasonably tight upper and lower bounds on this resolution along with the test underlying the upper bound.

3 Basic Test and Upper Resolution Bounds

In this section, we present a simple test which provides some upper bounds on the α-resolutions in problems (P_1) – (N_2).

Preliminaries. Let $\Gamma_N = \{\mu_\tau = \exp\{2\pi\imath\tau/N\} : 0 \leq \tau < N-1\}$ be the set of roots of 1 of degree N, and let $F_N : \mathbf{C}^N \to \mathbf{C}(\Gamma_N)$ be the normalized Fourier transform:

$$[F_N f](\mu) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} f_t \mu^t,\ \mu \in \Gamma_N. \tag{3}$$

Note that (3) can also be seen as a mapping from $\mathcal{S}$ to $\mathbf{C}(\Gamma_N)$.

Given a tolerance $\alpha \in (0, 1/2)$, let $q_N(\alpha)$ be the $(1-\alpha)$-quantile of $\|F_N\xi\|_\infty$, so that

$$\mathrm{Prob}_{\xi\sim\mathcal{N}(0,I_N)}\{\|F_N\xi\|_\infty \geq q_N(\alpha)\} \leq \alpha.$$

In the sequel we use the following immediate bound for $q_N(\cdot)$:

$$\begin{array}{rcl} q_N(\alpha) & \leq & \left\{\begin{array}{ll} \inf_{0\leq s\leq 1} \max\left[\mathrm{ErfInv}\left(\frac{s\alpha}{4}\right), \sqrt{\frac{N-2}{2(1-s)\alpha}}\right], & \text{for } N \text{ even}, \\ \inf_{0\leq s\leq 1} \max\left[\mathrm{ErfInv}\left(\frac{s\alpha}{2}\right), \sqrt{\frac{N-1}{2(1-s)\alpha}}\right], & \text{for } N \text{ odd} \end{array}\right. \\ & \asymp & \sqrt{\ln(N/\alpha)} \end{array} \tag{4}$$

Here $a \asymp b$ means that the ratio a/b is in-between positive absolute constants, and $\mathrm{ErfInv}(\alpha)$ is the inverse error function: $\frac{1}{\sqrt{2\pi}}\int_{\mathrm{ErfInv}(\alpha)}^{\infty} e^{-s^2/2}ds = \alpha$.

The test we are about to consider (and which we refer to as *basic test* in the sequel) is as follows:

1. given y, we solve the convex optimization problem

$$\mathrm{Opt}_{\mathcal{Z}}(y) = \min_{z\in\mathcal{Z}} \|F_N(y - z_0^{N-1})\|_\infty, \tag{5}$$

   where the set $\mathcal{Z}$ is defined according to

$$\mathcal{Z} = \left\{\begin{array}{ll} \{0\}, & \text{for problem } (P_1), \\ \mathcal{S}[\overline{\mathbf{w}}], & \text{for problem } (P_2), \\ \mathcal{S}^\epsilon[\overline{\mathbf{w}}], & \text{for problems } (N_1) \text{ and } (N_2). \end{array}\right. \tag{6}$$

2. We compare $\mathrm{Opt}_{\mathcal{Z}}(y)$ to $q_N(\alpha)$, where α is a given tolerance: if $\mathrm{Opt} \leq q_N(\alpha)$, we accept H_0, otherwise we accept H_1.

We describe now the properties of the basic test as applied to problems (P_1), (P_2), (N_1) and (N_2).

Theorem 3.1 *The risk of the basic test as applied to problems (P_1), (P_2) is bounded by α, provided that $d_* = d_n + d_s > 0$ and*

$$\rho \geq O(1)d_*^3 \ln(2d_*) q_N(\alpha) N^{-1/2} = O(1)d_*^3 \ln(2d_*)\sqrt{N^{-1}\ln(N/\alpha)} \tag{7}$$

with properly chosen positive absolute constants $O(1)$.

The result for the nonparametric problems (N_1) and (N_2) is similar:

Theorem 3.2 *The risk of the basic test as applied to problems* (N_1), (N_2) *is bounded by* α, *provided that* $d_* = d_n + d_s > 0$, ρ *satisfies (7) with properly selected* $O(1)$*'s and, in addition,* ϵ_n *and* ϵ_s *are small enough, specifically,*

$$N^{d_n+\frac{1}{2}}\epsilon_n + N^{d_s+\frac{1}{2}}\epsilon_s \leq O(1)q_N(\alpha) \tag{8}$$

with properly selected positive absolute constant $O(1)$.

The proofs of Theorems 3.1, 3.2 are relegated to section 6.

Theorems 3.1 and 3.2 provide us with upper resolution bounds independent of the frequencies constituting $\overline{\mathbf{w}}$ and $\mathbf{w}$. When ϵ_n, ϵ_s are "small enough," so that (8) holds true (we refer to the corresponding range of problems' parameters as the *parametric zone*), our upper bound on α-resolution in all testing problems of interest is essentially the same as in the case of $\epsilon_n = \epsilon_s = 0$ — it is $C(d_n + d_s)\sqrt{\ln(N/\alpha)/N}$ with the factor $C(d) = O(1)d^3\ln(2d)$ depending solely on d.

4 Lower Resolution Bounds

The lower resolution bounds of this section complement the upper bounds of section 3. We start with the parametric setting (P_1) and (P_2). Through this section, $c_i(d_n, d_s)$ are properly selected positive and monotone functions of their arguments.

Proposition 4.1 *Given integers* $d_n \geq 0$, $d_s \geq 1$, *and s real* $\alpha \in (0, 1/2)$, *consider problems* (P_1) *and* (P_2) *with the data* d_n ($d_n = 0$ *in the case of problem* (P_1)), d_s, α *and* $\overline{\mathbf{w}} = \{\overbrace{0, ..., 0}^{d_n}\}$. *Then for properly selected* $c_0(d_n, d_s)$ *and all* $N \geq c_0(d_n, d_s)$ *the* α*-resolution* ρ *in the problems* (P_1) *and* (P_2) *associated with the outlined data admits the lower bound*

$$O(1)d_s\sqrt{\ln(1/\alpha)/N}.$$

We see that in the problem (P_1) α-resolution grows with d_s at least linearly. Note that by Theorem 3.1, this growth is at most cubic (more precisely, it is not faster than $O(1)d_s^3\ln(d_s)$). Beside this, we see that the upper bounds on α-resolution for problems (P_1) and (P_2) stemming from Theorem 3.1 and associated with the basic test coincide, within a factor depending solely on d_n, d_s, N and logarithmic in N, with lower bounds on α-resolution.

We have the following lower bound on the α-resolution in the problem (N_1).

Proposition 4.2 *Given integers* $d_n > 0$, $d_s \geq 2$ *and reals* $\alpha \in (0, 1/2)$, $\epsilon_n \geq 0$, *consider problem* (N_1) *with the data* d_n, d_s, N, ϵ_n, α *and* $\overline{\mathbf{w}} = \{\overbrace{0, ..., 0}^{d_n}\}$. *Then for properly selected* $c_i(d_n, d_s) > 0$ *depending solely on* d_n, d_s *and for all* N *satisfying*

$$N \geq c_0(d_n, d_s), \tag{9}$$

the α*-resolution* $\rho_*(\alpha)$ *in the problem* (N_1) *associated with the outlined data satisfies:*

(i) in the range $0 \leq \epsilon_n \leq c_1(d_n, d_s) N^{-d_n - 1/2} \sqrt{\ln(1/\alpha)}$,

$$\rho_*(\alpha) \geq c_2(d_n, d_s) \sqrt{\ln(1/\alpha)/N};$$

(ii) in the range

$$c_3(d_n, d_s) N^{-d_n - 1/2} \sqrt{\ln(1/\alpha)} \leq \epsilon_n \leq c_1(d_n, d_s) N \sqrt{\ln(1/\alpha)}, \tag{10}$$

$$\rho_*(\alpha) \geq c_4(d_n, d_s) \left[\epsilon_n N^{d_n + 1/2} [\ln(1/\alpha)]^{-1/2}\right]^{\frac{1}{2d_n+3}} \sqrt{N^{-1} \ln(1/\alpha)}.$$

In the case of the problem (N_2) we have a similar lower bound on α-resolution when $\epsilon_n \leq \epsilon_s$.

Proposition 4.3 *Given integers* $d_n > 0$, $d_s \geq 2$ *and reals* $\alpha \in (0, 1/2)$, $\epsilon_n \geq 0$, *consider problem* (N_2) *with the data* d_n, d_s, N, ϵ_n, ϵ_s, α *and* $\overline{\mathbf{w}} = \{\overbrace{0, ..., 0}^{d_n}\}$.

Assume that $0 \leq \epsilon_n \leq \epsilon_s$. *Then for properly selected* $c_i(d_n, d_s) > 0$ *depending solely on* d_n, d_s *and all* $N \geq c_0(d_n, d_s)$ *the* α*-resolution* $\rho_*(\alpha)$ *in the problem* (N_2) *associated with the outlined data satisfies:*

(i) in the range $0 \leq \epsilon_s \leq c_1(d_n, d_s) N^{-d_s - 1/2} \sqrt{\ln(1/\alpha)}$,

$$\rho_*(\alpha) \geq c_2(d_n, d_s) \sqrt{\ln(1/\alpha)/N};$$

(ii) in the range

$$c_3(d_n, d_s) N^{-d_s - 1/2} \sqrt{\ln(1/\alpha)} \leq \epsilon_s \leq c_1(d_n, d_s) \sqrt{\ln(1/\alpha)}, \tag{11}$$

$$\begin{array}{rcl} \rho_*(\alpha) & \geq & c_4(d_n, d_s) \left[\epsilon_s N^{d_s + 1/2} [\ln(1/\alpha)]^{-1/2}\right]^{\frac{1}{2d_s+1}} \sqrt{N^{-1} \ln(1/\alpha)} \\ & \geq & c_5(d_n, d_s)\, \epsilon_s^{\frac{1}{2d_s+1}} (\ln(1/\alpha))^{\frac{d_s}{2d_s+1}}. \end{array} \tag{12}$$

The results of items (i) in Propositions 4.2 and 4.3 say that when d_n, d_s are fixed, N is large and ϵ_n, ϵ_s are small enough so that the problem parameters are in the parametric zone (i.e., (8) holds), Theorem 3.2 describes "nearly correctly" (i.e., up to factors depending solely on d_n, d_s, N and logarithmic in N) the α-resolution in problems (N_1) and (N_2): within such a factor, the α-resolution for problems (N_1), (N_2), same as for problems (P_1), (P_2), is $\sqrt{\ln(1/\alpha)/N}$. Besides, items (ii) in Propositions 4.2 and 4.3 show that when (ϵ_n, ϵ_s) goes "far beyond" the range (8), the α-resolution in problems (N_1), (N_2) becomes "much worse" than $\sqrt{\ln(1/\alpha)/N}$.

5 Numerical Results

Below we report on preliminary numerical experiments with the basic test.

5.1 Problem (N_1)

The goal of the first series of experiments was to quantify "practical performance" of the basic test as applied to problem (N_1).

Organization of experiments. We consider problem (N_1) on time horizon $0 \le t < N$ for $N \in \{128, 512, 1024\}$ with reliability threshold $\alpha = 0.01$. In these experiments $d_n = 4$, the frequencies in $\overline{\mathbf{w}}$ are selected at random, $\epsilon_n = 0.01$, and $d_s = 4$ (note that N and ϵ_n are deliberately chosen *not* to satisfy (8)). As explained in section 3, the above setup data specify the basic test for problem (N_1), and our goal is to find the "empirical resolution" of this test. To this end, we ran 10 experiments as follows. In a particular experiment,

- We draw at random $\mathbf{w} \in \mathcal{S}_4$, a *shift* $\bar{s} \in \mathcal{S}[\mathbf{w}]$ and *basic nuisance* $u \in \mathcal{S}^{\epsilon_n}[\overline{\mathbf{w}}]$.
- We generate a "true signal" x according to $x_\lambda = \lambda\bar{s} + u$, where $\lambda > 0$ is (nearly) as small as possible under the restriction that with $x = x_\lambda$, the basic test "rejects reliably" the hypothesis H_0, namely, rejects it in every one of 15 trials with $x = x_\lambda$ and different realizations of the observation noise ξ, see (1)[3].
- For the resulting λ, we compute $\rho = \min\limits_{u \in \mathcal{S}^{\epsilon_n}[\overline{\mathbf{w}}]} \|x_\lambda - u\|_\infty$, which is the output of the experiment. We believe that the collection of 10 outputs of this type gives a good impression on the "true resolution" of the basic test. As a byproduct of an experiment, we get also the $\|\cdot\|_\infty$-closest to x_λ point $u_x \in \mathcal{S}^{\epsilon_n}[\overline{\mathbf{w}}]$; the quantity $r = \|x_\lambda - u_x\|_2/\sqrt{N}$ can be thought of as a natural in our context "signal-to-noise ratio."

The results are presented in table 1. We would qualify them as quite compatible with the theory we have developed: both empirical resolution and empirical signal-to-noise ratio decreases with N as $N^{-1/2}$. The "empirically observed" resolution ρ for which the basic test $(1-\alpha)$-reliably, $\alpha = 0.01$, distinguishes between the hypotheses H_0 and $H_1(\rho)$ associated with problem (N_1) is $\approx 6\sqrt{\ln(N/\alpha)/N}$.

Comparison with MUSIC. An evident alternative to the basic test is (a) to apply the standard MUSIC algorithm [11] in order to recover the spectrum of the observed signal, (b) to delete from this spectrum the "nuisance frequencies", and (c) to decide from the remaining data if the signal of interest is present. Our related numerical results are, to the best of our understanding, strongly in favor of the basic test. Let us look at figure 1 where we present four MUSIC pseudospectra (we use `pmusic` function from MATLAB Signal Processing Toolbox) of the observations associated with signals x obeying the hypothesis $H_1(\rho)$ (magenta) and of the observations coming from the $\|\cdot\|_\infty$-closest to x nuisance (i.e., obeying the hypotheses H_0) u_x (blue). ρ was chosen large enough for the basic test to accept reliably the hypothesis $H_1(\rho)$ when it is true. We see that while sometimes MUSIC

[3] Since a run of the test requires solving a nontrivial convex program, it would be too time-consuming to replace 15 trials with few hundreds of them required to check reliably that the probability to reject H_0, the signal being x_λ, is at least the desired $1-\alpha = 0.99$.

Experiments with $N = 128$:

Experiment #	1	2	3	4	5	6	7	8	9	10	Mean	Mean$\times N^{1/2}$
Resolution	1.10	1.58	1.52	1.51	2.28	1.85	1.12	1.92	1.12	1.82	1.58	17.9
Signal/noise	0.70	1.09	0.94	0.95	1.36	1.03	0.77	1.06	0.69	1.10	0.97	11.0

Experiments with $N = 512$:

Experiment #	1	2	3	4	5	6	7	8	9	10	Mean	Mean$\times N^{1/2}$
Resolution	0.79	1.30	1.31	0.79	0.79	0.79	0.48	0.79	0.81	0.48	0.83	18.8
Signal/noise	0.44	0.71	0.74	0.43	0.43	0.44	0.30	0.50	0.44	0.34	0.48	10.8

Experiments with $N = 1024$:

Experiment #	1	2	3	4	5	6	7	8	9	10	Mean	Mean$\times N^{1/2}$
Resolution	0.60	0.92	0.36	0.58	0.46	0.35	0.36	0.56	0.42	0.59	0.52	16.6
Signal/noise	0.32	0.56	0.24	0.41	0.31	0.27	0.26	0.43	0.26	0.39	0.35	11.0

Table 1: Problem (N_1) with $d = d_s = 4$, $\alpha = \epsilon = 0.01$.

pseudospectrum indeed allows to understand which one of the hypotheses takes place (as it is the case in the example (d)), "MUSIC abilities" in our context are rather limited[4]. For example, it is hard to imagine a routine which would attribute magenta curves in the examples (a-c) to signals, and the blue curves — to the nuisances.

5.2 Comparison with Energy Test

Our objective here is to compare the resolution of the basic test to that of the test which implements the straightforward idea of how to discover if the signal x underlying observations (1) does not belong to a given nuisance set $\mathcal{U} \subset \mathcal{S}$. The test in question, which we refer to as *energy test*, is as follows: given a tolerance α and an observation y, we solve the optimization problem

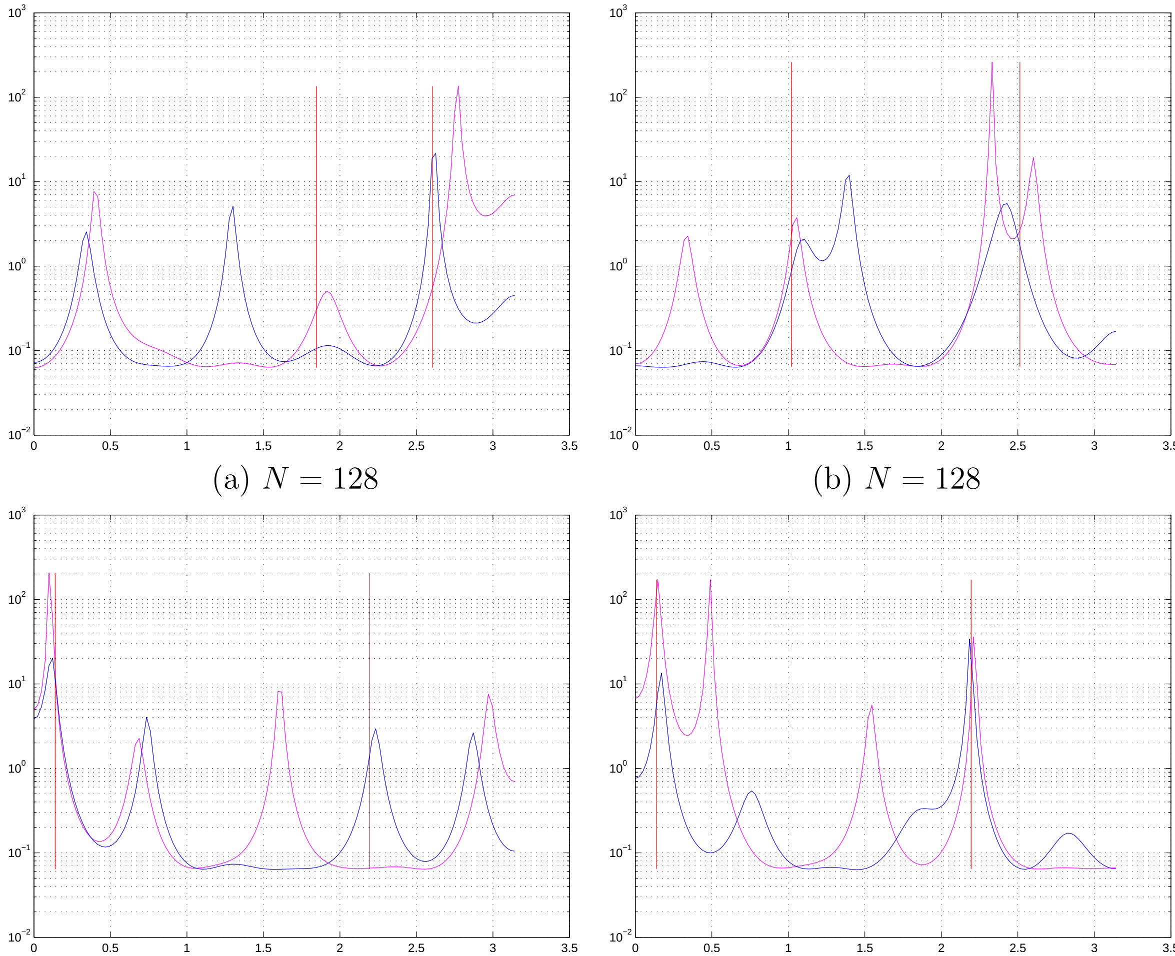
$$\mathrm{Opt}(y) = \inf_{u \in \mathcal{U}} \|y - u_0^{N-1}\|_2^2$$

and compare the optimal value with the $(1-\alpha)$-quantile

$$p_N(\alpha) : \mathrm{Prob}_{\xi \sim \mathcal{N}(0, I_N)}\{\|\xi\|_2^2 > p_N(\alpha)\} = \alpha$$

of the χ^2-distribution with N degrees of freedom. If $\mathrm{Opt}(y) > p_N(\alpha)$, we reject the nuisance hypothesis H_0 stating that $x \in \mathcal{U}$, otherwise we accept the hypothesis. Note that the basic test is of a completely similar structure, with $\|F_N(y - u_0^{N-1})\|_\infty^2$ in the role of $\|y - u_0^{N-1}\|_2^2$ and $q_N^2(\alpha)$ in the role of $p_N(\alpha)$. It is clear that the energy test rejects H_0 when the hypothesis is true with probability at most α (cf. item 1^0 in section 6.2). In order to reduce the computational burden, we restrict this test comparison to the simplest

[4]It should be noted that MUSIC is designed for a problem different from (and more complex than) the detection we are interested in, and thus its weakness relative to a dedicated detection test does not harm algorithm's well-established reputation.



(a) $N = 128$ (b) $N = 128$

(c) $N = 1024$ (d) $N = 1024$

Figure 1: MUSIC pseudospectra as built by MATLAB function `pmusic`$(\cdot,8)^\dagger$. Magenta: Signal plus nuisance; Blue: pure nuisance; Red bars: nuisance frequencies ($d = 4$ elements in $\overline{\mathbf{w}}$ correspond to 2 "actual" frequencies).

†In the present setup, $\overline{\mathbf{w}} \cup \mathbf{w} = \{\pm w_j, 1 \leq j \leq 4\}$, which requires the `pmusic` parameter `p` to be set to 8.

case of $\mathcal{U}=\{0\}$, i.e., the case of problem (P_1). Let us start with some theoretical analysis. Given a natural $d_s>0$ and a real $\rho>0$, consider the signal hypothesis $\widehat{H}_1(\rho)$ stating that the signal x underlying observations (1) satisfies $\|x_0^{N-1}\|_\infty\geq\rho$ and that x_t is a real algebraic polynomial of degree $\leq d-1$ of $t\in\mathbf{Z}$, meaning that $x\in\mathcal{S}[\overbrace{0,...,0}^{d}]$. Observe that with our $\mathcal{U}=\{0\}$, $\mathrm{Opt}(y)$ is nothing but

$$\|y\|_2^2=\|x_0^{N-1}+\xi\|_2^2=\|\xi\|_2^2+2\xi^Tx_0^{N-1}+\|x_0^{N-1}\|_2^2.$$

It follows that the hypothesis H_0 is accepted whenever the event

$$\|\xi\|_2^2+2\xi^Tx_0^{N-1}+\|x_0^{N-1}\|_2^2\leq p_N(\alpha)$$

takes place. Now, from the standard results on the χ^2 distribution it follows that for every $\alpha\in(0,1)$, for all large enough values of N with properly chosen absolute constants it holds

$$\mathrm{Prob}_{\xi\sim\mathcal{N}(0,I_N)}\{p_N(\alpha)-\|\xi\|_2^2\geq O(1)\sqrt{N\ln(1/\alpha)}\}\geq O(1),$$

whence also

$$\mathrm{Prob}_{\xi\sim\mathcal{N}(0,I_N)}\left\{\{p_N(\alpha)-\|\xi\|_2^2\geq O(1)\sqrt{N\ln(1/\alpha)}\}\cup\{\xi^Tx_0^{N-1}\leq 0\}\right\}\geq O(1).$$

As a result, whenever $x\in\mathcal{S}$ satisfies $\|x_0^{N-1}\|_2^2\leq O(1)\sqrt{N\ln(1/\alpha)}$, the probability to accept H_0, the true signal being x, is at least $O(1)$, provided that N is large enough. On the other hand, for a given d_s and large N there exists a polynomial x of degree d_s-1 such that $\|x_0^{N-1}\|_2\leq d_s^{-1}N^{-1/2}\|x_0^{N-1}\|_\infty$, see the proof of Proposition 4.1. It immediately follows that with $d_s\geq 1$ and (small enough) $\alpha>0$ fixed, the energy test cannot distinguish $(1-\alpha)$-reliably between the hypotheses H_0 and $\widehat{H}_1(\rho)$, provided that

$$\rho=O(1)d_s[N^{-1}\ln(1/\alpha)]^{1/4}\tag{13}$$

and N is large enough. In other words, with d_s and (small enough) α fixed, the resolution of the energy test in problem (P_1) admits, for large N, the lower bound (13). Note that as N grows, this bound goes to 0 as $N^{-1/4}$, while the resolution of the basic test goes to 0 as $N^{-1/2}\sqrt{\ln(N)}$ (Theorem 3.1). We conclue that the basic test provably outperforms the energy test as $N\to\infty$. The goal of the experiments we are about to report was to investigate this phenomenon numerically.

Organization of experiments. In the experiments to follow, the basic test and the energy test were tuned to 0.99-reliability ($\alpha=0.01$)) and used on time horizons $N\in\{256,1024,4096\}$. For a fixed N, and every value of the "resolution parameter" ρ from the equidistant grid on $[0,4]$ with resolution 0.05, we ran 10,000 experiments as follows:

- we generate $z\in\mathcal{S}[\mathbf{w}]$, and specify signal x as $\rho z/\|z_0^{N-1}\|_\infty$;
- we generate y according to (1) and run on the observations y the basic test and the energy test.

For every one of the tests, the outcome of a series of 10,000 experiments is the empirical probability p of rejecting the nuisance hypothesis H_0 (which states that the signal underlying the observations is identically zero). For $\rho = 0$, p is the (empirical) probability of false alarm (rejecting H_0 when it is true), and we want it to be small (about $\alpha = 0.01$). For $\rho > 0$, p is the empirical probability of successful detection of an actually present signal, and we want it to be close to 1 (about $1 - \alpha = 0.99$). Given that $p \leq \alpha$ when $\rho = 0$, the performance of a test can be quantified as the smallest value ρ_* of ρ for which p is at least $1 - \alpha$ (the less is ρ_*, the better).

We used 4-element collections $\mathbf{w}$ (i.e., used $d_s = 4$), and for every N and ρ ran two 10,000-element series of experiments differing in how we select $\mathbf{w}$ and z. In the first series ("random signals"), $\mathbf{w}$ was selected at random, and z was a random combination of the corresponding harmonic oscillations. In the second series ("bad signal") we used $\mathbf{w} = \{0, 0, 0, 0\}$, and z was the algebraic polynomial of degree 3 with the largest, among these polynomials, ratio of $\|z_0^{N_1}\|_\infty / \|z_0^{N-1}\|_2$. In the latter case only the realisation of noise varied from one experiment to another.

The results of our experiments are presented in table 2. They are in full accordance to what is suggested by our theoretical analysis; for $N = 256$, both tests exhibit nearly the same empirical performance. As N grows, the empirical performances of both tests improve, and the "performance gap" (which, as expected, is in favor of the basic test) grows.

6 Proofs

6.1 Preliminaries

Notation. In what follows, for a polynomial $p(\zeta) = \sum_{k=0}^m p_k \zeta^k$, we denote

$$\|p(\cdot)\|_\infty = \max_{\zeta \in \mathbf{C}, |\zeta|=1} |p(\zeta)|$$

and denote by

$$|p|_s = \|[p_0; p_1; ...; p_m]\|_s, \; 1 \leq s \leq \infty,$$

the ℓ_s-norm of the vector of coefficients, so that

$$\|p(\cdot)\|_2^2 := \frac{1}{2\pi} \oint_{|\zeta|=1} |p(\zeta)|^2 |d\zeta| = |p|_2^2.$$

The key fact underlying Theorems 3.1, 3.2 is the following

Proposition 6.1 *Let d, N be positive integers and $s \in \mathcal{S}_d$. Then*

$$\|F_N s\| \geq c(d) N^{1/2} \|s_0^{N-1}\|_\infty, \tag{14}$$

where $c(d) > 0$ is a universal nonincreasing function of d. One can take

$$c(d) = O(1)/(d^3 \ln(2d)) \tag{15}$$

with properly selected positive absolute constant $O(1)$.

$N = 256$, random signals ($\rho_*(\mathrm{B}) \approx 1.10, \rho_*(\mathrm{E}) \approx 1.35$) :

ρ	0.00	0.95	1.00	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40
p(B)	0.010	0.960	0.977	0.987	0.993	0.998	0.999	0.999	1.000	1.000	1.000
p(E)	0.011	0.710	0.779	0.842	0.887	0.933	0.956	0.974	0.987	0.994	0.997

ρ	1.45	1.50	1.55
p(B)	1.000	1.000	1.000
p(E)	0.999	1.000	1.000

$N = 256$, "bad" signal ($\rho_*(\mathrm{B}) \approx 2.65, \rho_*(\mathrm{E}) \approx 2.75$):

ρ	0.00	2.50	2.55	2.60	2.65	2.70	2.75	2.80	2.85	2.90	2.95
p(B)	0.010	0.973	0.984	0.987	0.991	0.994	0.998	0.998	0.999	1.000	1.000
p(E)	0.011	0.941	0.956	0.971	0.978	0.984	0.990	0.993	0.995	0.997	0.998

ρ	3.00	3.05	3.10
p(B)	1.000	1.000	1.000
p(E)	0.998	1.000	1.000

$N = 1024$, random signals ($\rho_*(\mathrm{B}) \approx 0.60, \rho_*(\mathrm{E}) \approx 0.90$):

ρ	0.00	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
p(B)	0.010	0.960	0.986	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p(E)	0.010	0.303	0.421	0.559	0.686	0.795	0.886	0.938	0.974	0.992	0.997

ρ	1.00	1.05	1.10
p(B)	1.000	1.000	1.000
p(E)	0.999	1.000	1.000

$N = 1024$, "bad" signal ($\rho_*(\mathrm{B}) \approx 1.40, \rho_*(\mathrm{E}) \approx 1.90$):

ρ	0.00	1.30	1.35	1.40	1.45	1.50	1.55	1.60	1.65	1.70	1.75
p(B)	0.011	0.960	0.980	0.990	0.997	0.998	0.999	1.000	1.000	1.000	1.000
p(E)	0.011	0.505	0.564	0.633	0.703	0.770	0.819	0.863	0.903	0.932	0.960

ρ	1.80	1.85	1.90	1.95	2.00
p(B)	1.000	1.000	1.000	1.000	1.000
p(E)	0.971	0.987	0.993	0.996	0.997

$N = 4096$, random signals ($\rho_*(\mathrm{B}) \approx 0.30, \rho_*(\mathrm{E}) \approx 0.65$):

ρ	0.00	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70
p(B)	0.009	0.931	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p(E)	0.009	0.084	0.165	0.291	0.472	0.667	0.823	0.930	0.980	0.997	1.000

$N = 4096$, "bad" signal ($\rho_*(\mathrm{B}) \approx 0.75, \rho_*(\mathrm{E}) \approx 1.35$):

ρ	0.00	0.70	0.75	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.15
p(B)	0.010	0.975	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
p(E)	0.012	0.184	0.223	0.290	0.376	0.477	0.567	0.676	0.767	0.843	0.899

ρ	1.20	1.25	1.30	1.35	1.40	1.45
p(B)	1.000	1.000	1.000	1.000	1.000	1.000
p(E)	0.945	0.975	0.986	0.995	0.999	1.000

Table 2: Basic test vs. Energy test, problem (P_1) with $d_s = 4$. p(B), p(E): empirical probabilities, taken over 10,000 trials, of detecting signal using the basic test (B) and the energy test (E). $\rho_*(\cdot)$; the smallest ρ for which $p(\cdot) \geq 1 - \alpha = 0.99$.

Proof. Let us fix $s \in \mathcal{S}_d$; we intend to prove that s obeys (14). Let $\mathbf{u} = \{\omega_1, ..., \omega_d\}$ be a symmetric mod 2π collection such that $s \in \mathcal{S}[\mathbf{u}]$, and let

$$p_{\mathbf{u}}(\zeta) = \prod_{\ell=1}^{d} (1 - \exp\{\imath\omega_\ell\}\zeta),$$

so that $p_{\mathbf{u}}(\Delta)s \equiv 0$. Let, further, M be the index of the largest in magnitude of the reals $s_0, s_1, ..., s_{N-1}$, so that

$$|s_M| = \|s_0^{N-1}\|_\infty. \tag{16}$$

We can w.l.o.g. assume that $M \geq (N-1)/2$. Indeed, otherwise we could pass from s to the "reversed" sequence $s' \in \mathcal{S}_d$: $s'_t = s_{N-t-1}$, $t \in \mathbf{Z}$, which would not affect the validity of our target relation (14) and would convert $M < (N-1)/2$ into $M' = N-1-M \geq (N-1)/2$.

1^0 We need the following technical

Lemma 6.1 *Let d be a positive integer, and let $\mathbf{u} = \{\upsilon_1, ..., \upsilon_d\} \in \Omega_d$. For every integer m satisfying*

$$m \geq m(d) := d\mathrm{Ceil}\left(5d \max[2, \tfrac{1}{2}\ln(2d)]\right) \tag{17}$$

one can point out real polynomials $q(\zeta) = \sum_{j=1}^{m} q_j\zeta^j$ and $r(\zeta) = 1 + \sum_{j=1}^{m-d} r_j\zeta^j$ such that

$$1 - q(\zeta) = p_{\mathbf{u}}(\zeta)r(\zeta), \tag{18}$$

and

$$|q|_2 \leq C_1(d)/\sqrt{m} \text{ where } C_1(d) = 3\mathrm{e}d^{3/2}\sqrt{\ln(2d)}. \tag{19}$$

The proof of Lemma 6.1 is presented in the appendix.

2^0. The following statement is immediate:

Lemma 6.2 *Let $m \leq (N-1)/2$, and $g \in \mathbf{C}^N$ be such that $g_i = 0$ for $i > m$. Let $h \in \mathbf{C}^N$ be the discrete autoconvolution of g, i.e. the vector with entries $h_k = \sum_{0\leq i,j\leq m, i+j=k} g_i g_j$, $0 \leq k \leq 2m$ and with zero remaining $N-2m-1$ entries. Then*

$$\|F_N h\|_1 = \sqrt{N}\|g\|_2^2.$$

Proof. We have

$$\begin{array}{rcl}[F_N h](\mu) & = & N^{-1/2}\sum_{0\leq t\leq 2m}\left[\sum_{0\leq j,k\leq m,\, j+k=t} g_j g_k\right]\mu^t \\ & = & N^{-1/2}\sum_{0\leq t\leq 2m}\sum_{0\leq j,k\leq m, j+k=t}(g_j\mu^j)(g_k\mu^k) \\ & = & N^{1/2}\left[N^{-1/2}\sum_{j=0}^{m} g_j\mu^j\right]^2.\end{array}$$

Invoking the Parseval identity, we conclude that

$$\|F_N h\|_1 = N^{1/2}\|F_N g\|_2^2 = N^{1/2}\|g\|_2^2. \quad \square$$

$\mathbf{3^0.}$ Let
$$N > 60d^2\ln(2d), \tag{20}$$
and let
$$m = \text{Floor}\left(\frac{N-1}{4}\right). \tag{21}$$
Then (17) is satisfied, and, according to Lemma 6.1, there exists a polynomial $q(\zeta) = \sum_{j=1}^m q_j\zeta^j$ such that
$$1 - q(\zeta) = p_{\mathbf{u}}(\zeta)r(\zeta),\ \ |q|_2 \le C_1(d)/\sqrt{m},$$
with some polynomial r. Setting $q^+(\zeta) = q^2(\zeta) = \sum_{j=1}^{2m} q_j^+\zeta^j$, we get $q_k^+ = \sum_{1\le i,j\le m,\, i+j=k} q_iq_j$
$$(1 - q^+(\Delta))s = ((1+q(\Delta))(1-q(\Delta))s = ((1+q(\Delta))r(\Delta)p_{\mathbf{u}}(\Delta)s \equiv 0,$$
whence $s_M = \sum_{i=1}^{2m} s_{M-i}q_i^+$ (note that $M \ge (N-1)/2 \ge 2m$ by (21)). Let now $h \in \mathbf{R}^N$ be the vector with coordinates
$$h_i = \left\{\begin{array}{ll} q^+_{M-i}, & i = M-1, ..., M-2m,\\ 0, & \text{otherwise.}\end{array}\right.$$
Note that by Lemma 6.2 and due to $|q|_2 \le C_1(d)/\sqrt{m}$ one has
$$\|F_Nh\|_1 \le N^{1/2}|q|_2^2 \le C_1^2(d)\sqrt{N}/m.$$
We have
$$\|s_0^{N-1}\|_\infty = |s_M| = \left|\sum_{i=1}^{2m} q_i^+ s_{M-i}\right| = |\langle h, s_0^{N-1}\rangle| = |\langle F_Nh, F_Ns\rangle| \le \|F_Nh\|_1\|F_Ns\|_\infty,$$
where the last equality is given by the fact that F_N is unitary, whence
$$\|F_Ns\|_\infty \ge \|s_0^{N-1}\|_\infty/\|F_Nh\|_1 \ge \frac{m}{C_1^2(d)\sqrt{N}}\|s_0^{N-1}\|_\infty.$$
Invoking (21), (20), and (19), we see that for N satisfying (20) our target relation (14) indeed holds true, provided that
$$c(d) \le O(1)[d^3\ln(2d)]^{-1} \tag{22}$$
with properly selected positive absolute constant $O(1)$.

$\mathbf{4^0.}$ It remains to verify (14) when $N \le 60d^2\ln(2d)$. Since F_N is unitary, we have $\|s_0^{N-1}\|_\infty \le \|s_0^{N-1}\|_2 = \|F_Ns\|_2 \le \|F_Ns\|_\infty\sqrt{N}$, whence
$$\|F_Ns\|_\infty \ge N^{-1/2}\|s_0^{N-1}\|_\infty \ge N^{-1}[N^{1/2}\|s_0^{N-1}\|_\infty] \ge \frac{1}{60d^2\ln(2d)}[N^{1/2}\|s_0^{N-1}\|_\infty],$$
which completes the proof. □

6.2 Proof of Theorem 3.1

1^0. Let us prove the result for the basic test, let it be denoted $\widehat{\varphi}$, as applied to the problem (P_2); note that (P_1) is the particular case of (P_2) corresponding to $d_n = 0$.

We have to show that under the premise of the theorem $\varepsilon_0(\widehat{\varphi}; H_0) \leq \alpha$ and $\varepsilon_1(\widehat{\varphi}; H_1(\rho)) \leq \alpha$. The first bound is evident. Indeed, let $\Xi_\alpha = \{\xi : \|F_N\xi\|_\infty \leq q_N(\alpha)\}$, so that $\mathrm{Prob}_{\xi\sim\mathcal{N}(0,I_N)}\{\xi \in \Xi_\alpha\} \geq 1-\alpha$. Under the hypothesis H_0, the set $\mathcal{Z}$ from (6) contains the true signal x_0^{N-1}, so that the optimal value $\mathrm{Opt}_{\mathcal{Z}}(y)$ in (5) is at most $\|F_N\xi\|_\infty$. It follows that when $\xi \in \Xi_\alpha$ (which happens with probability $\geq 1-\alpha$) we have $\mathrm{Opt} \leq q_N(\alpha)$, and the basic test will therefore accept H_0. We conclude that $\varepsilon_0(\widehat{\varphi}; H_0) \leq \mathrm{Prob}\{\xi \notin \Xi_\alpha\} \leq \alpha$.

2^0. Now let $x \in H_1(\rho)$, i.e., $x = s + u$, where $s \in \mathcal{S}[\mathbf{w}]$ for some $\mathbf{w} \in \Omega_{d_s}$, $u \in \mathcal{S}[\overline{\mathbf{w}}]$, and

$$\|[x - z]_0^{N-1}\|_\infty \geq \rho \ \forall z \in \mathcal{S}[\widehat{\mathbf{w}}].$$

Let $z \in \mathcal{S}[\widehat{\mathbf{w}}]$, and let $s = x - z$. Then $s \in \mathcal{S}_{d_*}$, $d_* = d_n + d_s$, and $\|s_0^{N-1}\|_\infty \geq \rho$, whence, by Proposition 6.1,

$$\|F_N s\|_\infty \geq c(d_*)N^{1/2}\|s_0^{N-1}\|_\infty \geq c(d_*)N^{1/2}\rho.$$

It follows that the optimal value $\mathrm{Opt}(y)$ in (5) is at least $c(d_*)N^{1/2}\rho - \|F_N\xi\|_\infty$. Recalling the definition of $q_N(\alpha)$, we conclude that

$$\mathrm{Prob}\{\mathrm{Opt}(y) > c(d_*)N^{1/2}\rho - q_N(\alpha)\} \geq 1 - \alpha$$

as soon as

$$\rho > \frac{2}{c(d_*)} q_N(\alpha)\sqrt{N}, \tag{23}$$

and the probability to reject $H_1(\rho)$ when the hypothesis is true is $\leq \alpha$. Specifying $c_1(d_*)$ as, say, $2.1c(d_*)^{-1}$, we see that under the premise of Theorem 3.1 one has $\varepsilon_1(\widehat{\varphi}) \leq \alpha$. □

6.3 Proof of Theorem 3.2

1^0. We start with the following simple

Lemma 6.3 *Let d and N be positive integers, let $\epsilon \geq 0$, let $\mathbf{u} = \{\upsilon_1, ..., \upsilon_d\}$ be a symmetric mod2π d-element collection of reals. Whenever $s \in \mathcal{S}^{N,\epsilon}[\mathbf{u}]$, there exists a decomposition $w = s + z$ such that $s \in \mathcal{S}[\mathbf{u}]$ and*

$$\|z_0^{N-1}\|_\infty \leq N^d\epsilon. \tag{24}$$

Proof. Let $p_{\mathbf{u}}(\zeta) = \prod_{\ell=1}^d (1 - \exp\{\imath\upsilon_\ell\}\zeta)$ and $r = p_{\mathbf{u}}(\Delta)w$, so that $\|r_0^{N-1}\|_\infty \leq \epsilon$ due to $w \in \mathcal{S}^{N,\epsilon}[\mathbf{u}]$. Let, further, δ be the discrete convolution unit (i.e., $\delta \in \mathcal{S}$ is given by $\delta_0 = 1$, $\delta_t = 0$, $t \neq 0$). For $\ell = 1, ..., d$, let $\gamma^{(\ell)}$ be a two-sided complex-valued sequence obtained from the sequence $\{\exp\{\imath\upsilon_\ell t\}\}_{t\in\mathbf{Z}}$ by replacing the terms with negative indexes with zeros, and let r^+ be obtained by similar operation from the sequence r. Let us set

$$\chi = \gamma^{(1)} * \gamma^{(2)} * ... * \gamma^{(d)} * r^+,$$

where $*$ stands for discrete time convolution. It is immediately seen that χ is a real-valued two-sided sequence which vanishes for $t < 0$ and satisfies the finite-difference equation $p_{\mathbf{u}}(\Delta)\chi = r^+$ (due to the evident relation $(1 - \exp\{\imath v_k\}\Delta)\gamma^{(k)} = \delta$). It follows that $(p_{\mathbf{u}}(\Delta)(w - \chi))_t = 0$ for $t = 0, 1, \ldots$, which (along with the fact that all the roots of $p_{\mathbf{u}}(\zeta)$ are nonzero) implies that the sequence $s = w - \chi$ can be modified on the domain $t < 0$ so that $p_{\mathbf{u}}(\Delta)s \equiv 0$. Then $z = w - s$ coincides with χ on the domain $t \geq 0$, and $w = s + z$ with $s \in \mathcal{S}[\mathbf{u}]$ and $z_t = \chi_t$, $t = 0, 1, \ldots$. It remains to note that for two-sided complex-valued sequences μ, ν starting at $t = 0$ we clearly have $\|[\mu * \nu]_0^{N-1}\|_\infty \leq \|\mu_0^{N-1}\|_1 \|\nu_0^{N-1}\|_\infty$. Applying this rule recursively and taking into account that $\|[\gamma^{(\ell)}]_0^{N-1}\|_1 = N$, we get the recurrence

$$\|[\gamma^{(1)} * \ldots * \gamma^{(\ell+1)} * r^+]_0^{N-1}\|_\infty \leq N \|[\gamma^{(1)} * \ldots * \gamma^{(\ell)} * r^+]_0^{N-1}\|_\infty, \; \ell = 0, 1, \ldots, d-1,$$

whence $\|\chi_0^{N-1}\|_\infty \leq N^d \epsilon$. Since $\chi_t = z_t$ for $t = 0, 1, \ldots$, (24) follows. $\square$

2^0. We are ready to prove Theorem 3.2. It suffices to consider the case of problem (N_2) (problem (N_1) is the particular case of (N_2) corresponding to $\epsilon_s = 0$). The fact that for the basic test $\widehat{\varphi}$ one has $\varepsilon_0(\widehat{\widehat{\phi}}) \leq \alpha$ can be verified exactly as in the case of Theorem 3.1. Let us prove that under the premise of Theorem 3.2 we have $\varepsilon_1(\widehat{\varphi}) \leq \alpha$ as well. To this end let the signal x underlying (1) belong to $H_1(\rho)$, so that $x = r + u$ for some $u \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}]$ and some $r \in \mathcal{S}^{N,\epsilon_n}[\mathbf{w}]$ with d_n-element collection $\overline{\mathbf{w}}$ and d_s-element collection $\mathbf{w}$, both symmetric mod 2π. Let also $z \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}]$. Since $x \in H_1(\rho)$, we have

$$\|[x - z]_0^{N-1}\|_\infty \geq \rho. \tag{25}$$

Applying Lemma 6.3 to r, u, z, we get the decompositions

$$\begin{array}{rcl} x & = & s + s' + v' : s \in \mathcal{S}[\mathbf{w}],\, s' \in \mathcal{S}[\overline{\mathbf{w}}],\, \|[v']_0^{N-1}\|_\infty \leq N^{d_n}\epsilon_n + N^{d_s}\epsilon_s, \\ z & = & s'' + v'' : s'' \in \mathcal{S}[\overline{\mathbf{w}}],\, \|[v'']_0^{N-1}\|_\infty \leq N^{d_n}\epsilon_n, \\ \Rightarrow w & := & x - z = \bar{s} + \bar{v}, \\ & & \bar{s} = s + s' - s'' \in \mathcal{S}[\overline{\mathbf{w}} \cup \mathbf{w}], \\ & & \bar{v} = v' - v'',\, \|\bar{v}_0^{N-1}\|_\infty \leq \sigma := 2N^{d_n}\epsilon_n + N^{d_s}\epsilon_s. \end{array} \tag{26}$$

Now, (25) implies that $\|w_0^{N-1}\| \geq \rho$, whence, by (26),

$$\|\bar{s}_0^{N-1}\|_\infty \geq \widehat{\rho} := \rho - \sigma.$$

Assuming that $\widehat{\rho} > 0$, noting that $\bar{s} \in \mathcal{S}[\widetilde{\mathbf{w}}]$ for $(d_* = d_n + d_s)$-element symmetric mod 2π collection $\widetilde{\mathbf{w}}$ and invoking Proposition 6.1, we get

$$\|F_N \bar{s}\|_\infty \geq c(d_*) N^{1/2} \widehat{\rho}.$$

Taking into account that $\|F_N \bar{v}\|_\infty \leq \|\bar{v}_0^{N-1}\|_2 \leq N^{1/2} \|\bar{v}_0^{N-1}\|_\infty$ and (26), we get also $\|F_N \bar{v}\|_\infty \leq N^{1/2}\sigma$. Combining these observations, we get

$$\begin{array}{rcl} \|F_N[x - z]\|_\infty & = & \|F_N \bar{s} + F_N \bar{v}\|_\infty \geq \|F_N \bar{s}\|_\infty - \|F_N \bar{v}\|_\infty \geq c(d_*) N^{1/2} \widehat{\rho} - N^{1/2}\sigma \\ & = & c(d_*) N^{1/2} [\rho - \sigma] - N^{1/2}\sigma = \underbrace{c(d_*) N^{1/2} \left[\rho - [1 + c_*^{-1}(d_*)]\sigma\right]}_{\vartheta}. \end{array}$$

Since $z \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}]$ is arbitrary, we conclude that the optimal value $\mathrm{Opt}(y)$ in (5) is at least $\vartheta - \|F_N\xi\|_\infty$, so that

$$\mathrm{Prob}\{\mathrm{Opt}(y) > \vartheta - q_N(\alpha)\} \geq 1-\alpha. \tag{27}$$

It remains to note that with properly selected positive absolute constants $O(1)$'s in (7) and (8), these restrictions on ρ, ϵ_n, ϵ_s ensure that $\vartheta > 2q_N(\alpha)$ (see (26), (15)), and therefore (27) implies the desired bound $\varepsilon_1(\widehat{\varphi}) \leq \alpha$. □

6.4 Proof of Proposition 4.1

Here we prove the lower resolution bound for problem (P_2). The result of the proposition for the setting (P_1) may be obtained by an immediate modification of the proof below for $d_n = 0$, $p_{\overline{\mathbf{w}}}(\cdot) = 1$, and $z \equiv 0$. Below we use notation κ_i for positive absolute constants.
1^0. Note that for every integer $d_s > 0$ there exists a real polynomial q_{d_s} on $[0,1]$ of degree $d_s - 1$ such that $\int_0^1 q_{d_s}^2(r)dr = 1$ and $\max_{0\leq r\leq 1}|q_{d_s}(r)| = q_{d_s}(1) = d_s$.[5] Let $q = \left\{q_t = N^{-1/2}q_{d_s}(t/N)\right\}_{t=-\infty}^{+\infty}$, $\lambda > 0$, and consider the two-sided sequence

$$\bar{x} = \left\{\bar{x}_t = \theta_N\lambda(-1)^t q_t\right\}_{t=-\infty}^{+\infty}$$

where $\theta_N > 0$ is chosen in such a way that

$$\|\bar{x}_0^{N-1}\|_2 = \lambda$$

Note that θ_N given by this requirement does not depend on λ and that $\theta_N \to 1$ as $N \to \infty$ due to $\int_0^1 q_{d_s}^2(r)dr = 1$. We have

$$|\bar{x}_{N-1}|/\|\bar{x}_0^{N-1}\|_2 \to \frac{|q_{d_s}(1)|}{\sqrt{\int_0^1 q_{d_s}^2(r)dr}} = d_s \text{ as } N \to \infty. \tag{28}$$

The derivative $q_{d_s}'(r)$ of the polynomial q_{d_s} satisfies

$$\max_{0\leq r\leq 1}|q_{d_s}'(r)| \leq 2(d_s-1)^2 \max_{0\leq r\leq 1}|q_{d_s}(r)| \leq 2d_s^3$$

(the first inequality in this chain follows from brother Markov's inequality). We conclude that for properly selected $\kappa_1 \geq 1$ and all $N \geq \kappa_1 d_s^2(d_n+1)$ it holds $q_{d_s}(t/N) \geq d_s/2$ whenever $N - d_n - 1 \leq t \leq N-1$. Taking into account that $\theta_N \to 1$ as $N \to \infty$, it follows that for properly selected $c_0(d_s,d_n) \geq \kappa_1 d_s^2(d_n+1)$ and all $N \geq c_0(d_s,d_n)$ it holds

$$\begin{array}{c} N - d_n - 1 \leq t \leq N-1 \Rightarrow |\bar{x}_t| = (-1)^t\bar{x}_t, \\ \min_{N-d_n-1\leq t\leq N-1}|\bar{x}_t| \geq |\bar{x}_{N-1}|/2 \geq \kappa_2\theta_N\lambda N^{-1/2}\,d_s \geq \kappa_3\lambda N^{-1/2}d_s \end{array} \tag{29}$$

[5] In fact, $\kappa+1$ is the maximal ratio of the uniform, on $[0,1]$, norm of a polynomial of degree $\leq \kappa$ and the $L_2(0,1)$-norm of the polynomial.

Beside this, for every d_s and N, we clearly have $\bar{x} \in \mathcal{S}[\mathbf{w}]$ with $\mathbf{w} = \{\overbrace{\pi, ..., \pi}^{d_s}\}$ (indeed, for all $t \in \mathbf{Z}$, $((1+\Delta)^{d_s}\bar{x})_t = (-1)^t\theta_N\lambda((1-\Delta)^{d_s}q)_t = 0)$.

2^0. Let $\overline{\mathbf{w}} = \{\overbrace{0, ..., 0}^{d_n}\}$, with $p_{\overline{\mathbf{w}}}(\zeta) = (1-\zeta)^{d_n}$. Assuming $N \geq c_0(d_s, d_n)$, we have

$$|(p_{\overline{\mathbf{w}}}(\Delta)\bar{x})_{N-1}| \geq 2^{d_n} \min_{N-d_n-1 \leq t \leq N-1} |\bar{x}_t| \geq 2^{d_n}\kappa_3\lambda N^{-1/2}\, d_s,$$

so that for every $z \in \mathcal{S}[\overline{\mathbf{w}}]$ we have

$$|(p_{\overline{\mathbf{w}}}(\Delta)[\bar{x} - z])_{N-1}| \geq 2^{d_n}\kappa_3\lambda N^{-1/2}\, d_s.$$

Since $N - 1 \geq d_n$, when taking into account that $|p_{\overline{\mathbf{w}}}|_1 = 2^{d_n}$ we get for any $z \in \mathcal{S}[\overline{\mathbf{w}}]$

$$2^{d_n}\|[\bar{x} - z]_0^{N-1}\|_\infty = |p_{\overline{\mathbf{w}}}|_1\|[\bar{x} - z]_0^{N-1}\|_\infty \geq |(p_{\overline{\mathbf{w}}}(\Delta)[\bar{x} - z])_{N-1}| \geq 2^{d_n}\kappa_3\lambda N^{-1/2}\, d_s,$$

whence

$$\forall\, z \in \mathcal{S}[\widehat{\mathbf{w}}]: \;\; \|[\bar{x} - z]_0^{N-1}\|_\infty \geq \kappa_3\lambda N^{-1/2}\, d_s. \tag{30}$$

Now let us set $\lambda = \kappa_4\sqrt{\ln(1/\alpha)}$, with the absolute constant κ_4 selected to ensure that $\lambda < 2\mathrm{ErfInv}(\alpha)$. The latter relation, due to $\lambda = \|x_0^{N-1}\|_2$, ensures that the hypotheses *"observation (1) comes from $x \equiv 0$"* and *"observation (1) comes from $x = \bar{x}$"* cannot be distinguished $(1-\alpha)$-reliably. Thus, (30) implies the lower resolution bound $\kappa_3\kappa_4 d_s N^{-1/2}\sqrt{\ln(1/\alpha)}$. □

6.5 Proof of Proposition 4.2

In the reasoning below, c_i denote positive quantities depending solely on $d = d_n$, and κ_i denote positive absolute constants. We start with proving the claim (ii).

1^0. Let us set

$$\bar{f}_t = \sin\left(\frac{\pi}{8} + \frac{\pi}{4}\frac{t}{N-1}\right), \;\; t \in \mathbf{Z}. \tag{31}$$

Assuming $c_0 > 40d$, see (9.a), let us fix an integer τ such that

$$d - 1 \leq \tau \leq (N-1) - 20d \tag{32}$$

and set

$$\gamma = \gamma_{\tau,N} = \epsilon\bar{f}_\tau^{-1}, \;\; f = \gamma\bar{f}. \tag{33}$$

By the definition of f we have

$$0 \leq f_t \leq \epsilon \quad \text{for} \;\; 0 \leq t \leq \tau, \tag{34}$$

and

$$\epsilon + \kappa_0 N^{-1}(t-\tau)\epsilon \leq f_t \leq \epsilon + \kappa_1 N^{-1}(t-\tau)\epsilon \;\; \text{for} \;\; \tau \leq t \leq N-1. \tag{35}$$

Let $p(\zeta)=(1-\zeta)^d=p_{\overline{\mathbf{w}}}(\zeta)$. We clearly can find a sequence $x=\{x_t=a\cos(\frac{\pi}{4}t+b)\}_{t=-\infty}^{\infty}$ such that $p(\Delta)x=f$, and, due to $d_s\geq 2$, we have $x\in\mathcal{S}_{d_s}$. Further, let $\bar{z}\in\mathcal{S}$ satisfy $\bar{z}_t=x_t$ for $0\leq t\leq\tau$, and

$$(p(\Delta)\bar{z})_t=\begin{cases} f_t, & 0\leq t\leq\tau,\\ 0, & t<0,\\ \epsilon, & t\geq\tau.\end{cases}$$

Note that $\bar{z}$ is well defined due to $p(\Delta)x=f$, and, taking into account (33) – (35), we conclude that $\bar{z}\in\mathcal{S}^\epsilon[\overline{\mathbf{w}}]$.

2^0. By the above construction, the sequence $\delta=x-\bar{z}$ is such that $(p(\Delta)\delta)_t=0$ for $0\leq t\leq\tau$, and $\kappa_0(t-\tau)N^{-1}\epsilon\leq(p(\Delta)\delta)_t\leq\kappa_1(t-\tau)N^{-1}\epsilon$ when $\tau<t<N$, see (35). Besides, $\delta_t=0$ when $0\leq t\leq\tau$. By evident reasons, these two observations combine with the first inequality in (32) to imply that

$$\|\delta_0^{N-1}\|_\infty\leq c_4(N-\tau)^{d+1}N^{-1}\epsilon. \tag{36}$$

for some $c_4>0$ depending solely on d. We conclude that

$$\|\delta_0^{N-1}\|_2\leq\|\delta_0^{N-1}\|_\infty\sqrt{N-\tau}\leq c_4(N-\tau)^{d+3/2}N^{-1}\epsilon.$$

Now note that the hypotheses "observation (1) comes from $x=\bar{z}$" and "observation (1) comes from $x=\bar{z}+\delta$" cannot be distinguished $(1-\alpha)$-reliably unless $\|\delta_0^{N-1}\|_2\geq 2\mathrm{ErfInv}(\alpha)>\sqrt{\kappa_3\ln(1/\alpha)}$. Equipped with this κ_3 and with c_4 participating in (36), let us set

$$\nu=\mathrm{Floor}\left(\left(\kappa_3\ln(1/\alpha)N^2c_4^{-2}\epsilon^{-2}\right)^{\frac{1}{2d+3}}\right). \tag{37}$$

It is immediately seen that with properly chosen positive c_1,c_3 depending solely on d, and with ϵ satisfying (10), we have $20d<\nu<N-d$, so that setting

$$\tau=N-\nu,$$

we ensure (32). From now on, we assume that c_1,c_3 are as needed in the latter conclusion. With the just defined τ, we have $\|\delta_0^{N-1}\|_2^2\leq\kappa_3\ln(1/\alpha)$, meaning that x and $\bar{z}$ cannot be distinguished $(1-\alpha)$-reliably.

3^0. To prove the claim (ii) it now suffices to show that a properly chosen $c_5>0$, depending solely on d,

$$\forall z\in\mathcal{S}^\epsilon[\overline{\mathbf{w}}]:\|[x-z]_0^{N-1}\|_\infty\geq c_5(N-\tau)^{d+1}N^{-1}\epsilon=c_5\nu^{d+1}N^{-1}\epsilon. \tag{38}$$

Indeed, given $z\in\mathcal{S}^\epsilon$, we have

$$\tau\leq t<N\Rightarrow(p(\Delta)[x-z])_t=(p(\Delta)x)_t-(p(\Delta)z)_t\geq f_t-\epsilon\geq\kappa_0N^{-1}\epsilon(t-\tau),$$

with concluding inequality given by (35). Setting $\theta=\lfloor(N-1-\tau)/2\rfloor$, the sequence

$$s_t=(x-z)_{t-\tau},\quad t\in\mathbf{Z}$$

satisfies

$$(p(\Delta)s)_t \geq \kappa_2 \theta N^{-1} \epsilon, \;\; \theta \leq t \leq 2\theta, \tag{39}$$

and, by the right inequality in (32), $\theta \geq 10d$. Setting $k = \lfloor (\theta - d)/d \rfloor \geq 2$, consider the polynomial

$$q(\zeta) = (1 - \zeta^k)^d = (1 - \zeta)^d \underbrace{(1 + \zeta + ... + \zeta^{k-1})^d}_{v(\zeta)} = (1 - \zeta)^d \sum_{j=0}^{d(k-1)} v_j \zeta^j;$$

where, clearly,

$$v_j \geq 0, \;\; \sum_j v_j = k^d. \tag{40}$$

Let now $r = q(\Delta)s$. Taking into account that $|q(\cdot)|_1 = 2^d$, we have

$$r_{2\theta} \leq 2^d \|s_\theta^{2\theta}\|_\infty \leq 2^d \|[x - z]_0^{N-1}\|_\infty, \tag{41}$$

since, by construction, $\|[x - z]_0^{N-1}\|_\infty \geq \|s_\theta^{2\theta}\|_\infty$. On the other hand, $r = v(\Delta)u$, $u = p(\Delta)s$, so that

$$r_{2\theta} = \sum_{j=0}^{d(k-1)} v_j u_{2\theta - j}.$$

By (39), we have $u_{2\theta-j} \geq \kappa_2 \theta N^{-1} \epsilon$ for $0 \leq j \leq d(k-1)$ (note that $d(k-1) < \theta$), and by (40),

$$r_{2\theta} \geq \kappa_2 \theta N^{-1} \epsilon \sum_j v_j = k^d \kappa_2 \theta N^{-1} \epsilon.$$

Combining the latter inequality with (41) we come to $\|[\bar{x} - z]_0^{N-1}\|_\infty \geq \kappa_2 2^{-d} k^d \theta N^{-1} \epsilon$. Recalling that by construction $k > \kappa_{10}(N - \tau)/d$, $\theta \geq \kappa_{11}(N - \tau)$, we arrive at (38). (ii) is proved.

4^0. It remains to prove (i). Note that when $\epsilon < c_{12} N^{-1/2} \sqrt{\ln(1/\alpha)}$, the conclusion in (i) is readily given by a straightforward modification of the reasoning in section 6.4. Note that for the time being the only restriction on the lower bound c_0 on N, see (9.a), was that $c_0 \geq 40d$, see the beginning of item 1^0. Now let us also assume that N is large enough to ensure that $c_{12} N^{-1/2} \geq c_3 N^{-d-1/2}$ (which still allows to choose c_0 as a function of $d = d_n$ only). With the resulting c_0, the range of values of ϵ for which we have justified the conclusion in (i) covers the corresponding range of ϵ allowed by the premise of (i). □

6.6 Proof of Proposition 4.3

The proof of the statement (i) is given by a straightforward modification of the reasoning in section 6.4. Let us prove (ii). Let $\mathbf{w} = \{\overbrace{\pi, ..., \pi}^{d_s}\}$, and $0 \leq \tau < N - d_n - 1$, so that $\mathcal{S}^{N,\epsilon_s}[\mathbf{w}]$ contains the sequence $\bar{x} = \{\bar{x}_t = c_1 \epsilon_s (-1)^t (t - \tau)_+^{d_s}\}_{t=-\infty}^{\infty}$ with $0 < c_1 = c_1(d_s)$ small enough. Then for $\overline{\mathbf{w}} = \{0, ..., 0\}$ (d_n zeros), with $p_{\overline{\mathbf{w}}}(\zeta) = (1 - \zeta)^{d_n}$, we have for any $z \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}]$:

$$\begin{array}{rcl} |(p_{\overline{\mathbf{w}}}(\Delta)[\bar{x} - z])_{N-1}| & \geq & |(p_{\overline{\mathbf{w}}}(\Delta)\bar{x})_{N-1}| - |(p_{\overline{\mathbf{w}}}(\Delta)z)_{N-1}| \geq 2^{d_n} |\bar{x}_{N-d-1}| - \epsilon_n \\ & \geq & c_2 \epsilon_s (N - \tau)^{d_s} - \epsilon_n, \end{array}$$

where $c_2 > 0$ depends only on d_s and d_n. Therefore, when denoting $\nu = N - \tau$, for every $z \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}]$ we have

$$2^{d_n}\|[\bar{x} - z]_0^{N-1}\|_\infty = |p_{\overline{\mathbf{w}}}|_1\|[\bar{x} - z]_0^{N-1}\|_\infty \geq |(p_{\overline{\mathbf{w}}}(\Delta)[\bar{x} - z])_{N-1}| \geq c_2\epsilon_s\nu^{d_s} - \epsilon_n,$$

whence,

$$\forall\, z \in \mathcal{S}^{N,\epsilon_n}[\overline{\mathbf{w}}]: \quad \|[\bar{x} - z]_0^{N-1}\|_\infty \geq 2^{-d_n}(c_2\epsilon_s\nu^{d_s} - \epsilon_n). \tag{42}$$

It is immediately seen that with properly selected positive $c_i = c_i(d_n, d_s)$, $i = 3, 4, 5, 6$, assuming

$$\epsilon_s N^{d_s+\frac{1}{2}} \geq c_3\sqrt{\ln(1/\alpha)} \ \&\ \epsilon_s \leq c_4\sqrt{\ln(1/\alpha)} \tag{43}$$

and selecting τ according to

$$\nu = N - \tau = \mathrm{Floor}\left(c_5\left[\epsilon_s^{-2}\ln(1/\alpha)\right]^{\frac{1}{2d_s+1}}\right),$$

we ensure that $0 \leq \tau < N - d_n - 1$ and

$$\begin{array}{ll} (a) \quad \tau < N - d_n - 1, & (b) \quad \|\bar{x}_0^{N-1}\|_2 < 2\mathrm{ErfInv}(\alpha), \\ (c) \quad 2^{-d_n}(c_2\epsilon_s\nu^{d_s} - \epsilon_n) \geq \bar{\rho} := c_6\epsilon_s^{\frac{1}{2d_s+1}}(\ln(1/\alpha))^{\frac{d_s}{2d_s+1}}, & \end{array}$$

(when verifying (c), take into account that $0 \leq \epsilon_n \leq \epsilon_s$). By (b), the hypotheses "observation (1) comes from $x \equiv 0$" and "observation (1) comes from $x = \bar{x}$" cannot be distinguished $(1-\alpha)$-reliably, while (42), the already established inclusion $\bar{x} \in \mathcal{S}^{N,\epsilon_s}[\mathbf{w}]$ and (c) imply that $\bar{x}$ obeys the hypothesis $H_1(\bar{\rho})$ associated with the problem (N_2). The bottom line is that in the case of (43), $\bar{\rho}$ defined in (c) is a lower bound on $\rho_*(\alpha)$. $\square$

References

[1] Chiu, S.-T., "Detecting periodic components in a white noise gaussian time series", *J. R. Statist. Soc., Ser. B* **51**, (1989) 249-359.

[2] Goldenshluger, A., and Nemirovski, A., "Spatial adaptive estimation of functions satisfying differential inequalities" – *IEEE Transactions on Information Theory* **43** (1997), 872-889.

[3] Fisher, R. A., "Test of significance in harmionic analysis" – *Proc. Royal Soc. London, Ser. A* **125**, No. 796 (1929), 54-59.

[4] Hannan, E. J., *"Multiple Time Series"*, New York: Wiley (1970).

[5] Hannan, E. J., Quinn, B. G., *"The Estimation and Tracking of Frequency"*, Cambridge University Press (2001).

[6] Juditsky, A., Nemirovski, A., "Nonparametric denoising of signals with unknown local structure, I: Oracle inequalities" – *Applied and Computational Harmonic Analysis* **27** (2009), 157-179.

[7] Juditsky, A., Nemirovski, A., "Nonparametric denoising signals with unknown local structure, II: Nonparametric function recovery" – *Applied and Computational Harmonic Analysis*, **29:3** (2010) 354-367.

[8] Nemirovskii, A., "On forecast under uncertainty" (in Russian) – *Problemy Peredachi Informacii*, **17:4** (1981) (the journal is translated into English as *Problems of Information Transmission*)

[9] Nemirovskii, A., "Nonparametric estimation of functions satisfying differential inequalities" - in: R. Khasminskii, Ed. *Advances in Soviet Mathematics, v. 12: "Selected Topics in Nonparametric Statistics"* - American Mathematical Society, Providence, 1992

[10] Pisarenko, V. F., "The Retrieval of Harmonics from a Covariance Function" *Geophys. J. Roy. Astr. Soc.*, **33** (1973), 347-366.

[11] Schmidt, R.O., "Multiple emitter location and signal parameter estimation", *Proc. RADC, Spectral Estimation Workshop, Rome, NY* (1979), 243-258.

[12] Stoica, P., Moses, R.L., *Introduction to Spectral Analysis*, Prentice-Hall (1997).

[13] Whittle, P., The statistical analysis of a seiche record. *J. of Marine Res.*, **13** (1954), 76-100.

A Proof of Lemma 6.1

1^0. Let $\lambda \in \mathbf{C}$, $0 < |\lambda| \leq 1$, let $\epsilon \in (0,1)$, and let $n \geq 1$ be an integer. Setting $\delta = 1-\epsilon$, consider the polynomials

$$\begin{array}{rcl} f_k(\zeta) & = & f_k(z;\lambda,\epsilon) := (1-\lambda\zeta)\sum_{\ell=0}^k (\delta\lambda\zeta)^\ell = \frac{1-\lambda z}{1-\delta\lambda z}[1-[\delta\lambda\zeta]^{k+1}],\ k=0,1,...\\ p_n(\zeta) & = & p_n(z;\lambda,\epsilon) := \frac{1}{n}\sum_{k=0}^{n-1} f_k(\zeta) = (1-\lambda\zeta)\sum_{\ell=0}^{n-1}[\delta\lambda\zeta]^\ell[1-\ell/n] \end{array} \tag{44}$$

Observe that

$$f_k(0)=1,\, f_k(1/\lambda)=0,$$

whence

$$p_n(0)=1,\ p_n(1/\lambda)=0. \tag{45}$$

1.1^0. Let us bound from above the uniform norm $\|p_n(\cdot)\|_\infty$ of p_n on the unit circle. We have $p_n(\zeta) = r_n(\lambda\zeta)$, where

$$\begin{array}{rcl} r_n(\zeta) & = & \frac{1}{n}(1-\zeta)\sum_{k=0}^{n-1}\frac{1-(\delta\zeta)^{k+1}}{1-\delta\zeta} = (1-\zeta)\frac{n-\delta\zeta\frac{1-[\delta\zeta]^n}{1-\delta\zeta}}{n(1-\delta\zeta)}\\ & = & \underbrace{\frac{1-\zeta}{1-\delta\zeta}}_{g_n(\zeta)}\underbrace{\frac{n(1-\delta\zeta)-\delta\zeta+[\delta\zeta]^{n+1}}{n(1-\delta\zeta)}}_{h_n(\zeta)}. \end{array}$$

Since $|\lambda| \leq 1$, we have $\|p_n(\cdot)\|_\infty \leq \|g_n(\cdot)\|_\infty \|h_n(\cdot)\|_\infty$. When $|\zeta| = 1$, we have

$$|g_n(\zeta)| = \frac{1}{\delta} \frac{|1-\zeta|}{|\delta^{-1} - \zeta|} \leq \frac{1}{\delta}. \tag{46}$$

If we set $\zeta = \cos(\phi) + \imath \sin(\phi)$, and $\epsilon = \theta/n$ and $\delta = 1 - \epsilon = 1 - \theta/n$ with some $\theta \in (0, n)$, we obtain

$$\begin{aligned}
|h_n(\zeta)| &\leq \frac{|n(1-\delta\zeta) - \delta\zeta| + \delta^{n+1}}{n|1-\delta\lambda|} \\
&= \frac{\sqrt{[n-(n+1)\delta\cos(\phi)]^2 + (n+1)^2\delta^2\sin^2(\phi)} + \delta^{n+1}}{n\sqrt{[1-\delta\cos(\phi)]^2 + \delta^2\sin^2(\phi)}} \\
&= \frac{\sqrt{n^2 + (n+1)^2\delta^2\cos^2(\phi) - 2n(n+1)\delta\cos(\phi) + (n+1)^2\delta^2\sin^2(\phi)} + \delta^{n+1}}{n\sqrt{1 + \delta^2\cos^2(\phi) - 2\delta\cos(\phi) + \delta^2\sin^2(\phi)}} \\
&= \frac{\sqrt{n^2 + (n+1)^2 - 2n(n+1)\delta\cos(\phi)} + \delta^{n+1}}{n\sqrt{1 + \delta^2 - 2\delta\cos(\phi)}} \\
&= \frac{\sqrt{[n-(n+1)\delta)]^2 + 2n(n+1)\delta[1-\cos(\phi)]} + \delta^{n+1}}{n\sqrt{[1-\delta]^2 + 2\delta[1-\cos(\phi)]}} \\
&= \frac{\sqrt{\left[\theta\frac{n+1}{n} - 1\right]^2 + 2(n+1)(n-\theta)[1-\cos(\phi)]} + \delta^{n+1}}{n\sqrt{\frac{\theta^2}{n^2} + 2n^{-1}(n-\theta)[1-\cos(\phi)]}} \\
&= \frac{\sqrt{\left[\theta\frac{n+1}{n} - 1\right]^2 + 2(n+1)(n-\theta)[1-\cos(\phi)]} + \delta^{n+1}}{\sqrt{\theta^2 + 2n(n-\theta)[1-\cos(\phi)]}} = \frac{\sqrt{\alpha + \beta t} + \gamma}{\sqrt{\mu + \nu t}},
\end{aligned}$$

where

$$\begin{aligned}
&\alpha = \left[\theta\tfrac{n+1}{n} - 1\right]^2, \beta = 2(n+1)(n-\theta),\ \gamma = \delta^{n+1},\ \mu = \theta^2,\ \nu = 2n(n-\theta), \\
&t = 1 - \cos(\phi)
\end{aligned} \tag{47}$$

It is immediately seen that with positive $\alpha, \beta, \gamma, \mu$ and ν such that

$$\mu/\nu > \alpha/\beta, \tag{48}$$

the maximum of the function $\frac{\sqrt{\alpha+\beta t}+\gamma}{\sqrt{\mu+\nu t}}$ over t such that $\alpha + \beta t \geq 0$ is achieved when $\alpha + \beta t = \sqrt{\frac{\beta\mu - \alpha\nu}{\gamma\nu}}$ and is equal to $\sqrt{\frac{\beta}{\nu}}\sqrt{1 + \frac{\gamma^2\nu}{\beta\mu - \alpha\nu}}$. Now, assume that

$$1 < \theta < n. \tag{49}$$

Then $\frac{n+1}{n}\theta^2 > [\theta\frac{n+1}{n} - 1]^2$, so that the parameters $\alpha, ..., \nu$ defined in (47) satisfy (48). We conclude that

$$\|h_n(\cdot)\|_\infty \leq \sqrt{\frac{n+1}{n}} \sqrt{1 + \frac{[1-\theta/n]^{2(n+1)}}{\frac{n+1}{n}\theta^2 - \left[\theta\frac{n+1}{n} - 1\right]^2}}. \tag{50}$$

Note that
$$\frac{n+1}{n}\theta^2-[\theta\frac{n+1}{n}-1]^2=2\theta\frac{n+1}{n}-\theta^2\frac{n+1}{n^2}-1\geq 1$$
when θ satisfies
$$\frac{2n}{n+1}\leq\theta<n, \tag{51}$$
and (50) implies in this case that
$$\|h_n(\cdot)\|_\infty\leq\sqrt{\frac{n+1}{n}}\sqrt{1+[1-\theta/n]^{2(n+1)}}\leq\exp\left\{\frac{1}{2n}+\frac{1}{2}\mathrm{e}^{-2\theta}\right\}.$$
The latter bound combines with (46) to imply that in the case of (51) (recall that $1-\delta=\epsilon=\theta/n$) we get for all $|\lambda|\leq 1$:
$$\begin{array}{rcl}\max\limits_{|z|\leq 1}|p_n(z;\lambda,\epsilon)| & \leq & \frac{1}{1-\theta/n}\exp\left\{\frac{1}{2n}+\frac{1}{2}\mathrm{e}^{-2\theta}\right\}\\ & \leq & \exp\left\{\frac{\theta}{n-\theta}+\frac{1}{2n}+\frac{1}{2}\mathrm{e}^{-2\theta}\right\}.\end{array} \tag{52}$$

1.2^0. Now let us bound from above $\|1-p_n(\cdot)\|_2$. We have
$$\begin{array}{rcl}p_n(\zeta) & = & 1+\sum_{\ell=1}^{n-1}\left(\delta^\ell\left[1-\frac{\ell}{n}\right]-\delta^{\ell-1}\left[1-\frac{\ell-1}{n}\right]\right)[\lambda\zeta]^\ell-\frac{1}{n}\delta^{n-1}[\lambda\zeta]^n\\ & = & 1+\sum_{\ell=1}^{n-1}\delta^{\ell-1}\left([1-\epsilon]\left[1-\frac{\ell}{n}\right]-\left[1-\frac{\ell-1}{n}\right]\right)[\lambda\zeta]^\ell-\frac{1}{n}\delta^{n-1}[\lambda\zeta]^n\\ & = & 1-\left(\sum_{\ell=1}^{n-1}\delta^{\ell-1}\left[\epsilon[1-\frac{\ell}{n}]+\frac{1}{n}\right][\lambda\zeta]^\ell+\frac{1}{n}\delta^{n-1}[\lambda\zeta]^n\right)\end{array}$$
Taking into account that $|\lambda|\leq 1$, we conclude that
$$\begin{array}{rcl}\|1-p_n(\cdot)\|_2 & \leq & \|\sum\limits_{\ell=1}^{n-1}(1-\epsilon)^{\ell-1}\epsilon[1-\frac{\ell}{n}][\lambda\zeta]^\ell\|_2+\frac{1}{n}\|\sum\limits_{\ell=1}^{n}(1-\epsilon)^{\ell-1}[\lambda\zeta]^\ell\|_2\\ & \leq & \sqrt{\sum\limits_{\ell=1}^{n}\epsilon^2(1-\epsilon)^{2(\ell-1)}}+\sqrt{1/n}\\ & \leq & \sqrt{\frac{\epsilon^2}{1-(1-\epsilon)^2}}+\sqrt{1/n}\leq\sqrt{\epsilon}+\sqrt{\frac{1}{n}}\leq 2\sqrt{\max\left[\epsilon,\frac{1}{n}\right]}.\end{array} \tag{53}$$

2^0. Let $n=\lfloor m/d\rfloor$ and $\lambda_\ell=\exp\{-\imath v_\ell\}$, $1\leq\ell\leq d$, where m, d and $v_1,...,v_d$ are as described in the premise of the lemma. Let us set
$$\begin{array}{rcl}\theta & = & \max[2,\frac{1}{2}\ln(2d)]\leq 3\ln(2d),\ \epsilon=\frac{\theta}{n},\\ q(\zeta) & = & 1-p_n(z;\lambda_1,\epsilon)\cdot p_n(z;\lambda_2,\epsilon)\cdot\ldots\cdot p_n(\lambda_d,\epsilon).\end{array}$$

2.1^0. Observe that by (17) we have

$$n \geq n(d) := \mathrm{Ceil}(5d\theta) = \mathrm{Ceil}(5d\max[2, \frac{1}{2}\ln(2d)]). \tag{54}$$

Our choice of θ and n ensures (51), so that $0 < \epsilon < 1$, and by (52) we also have

$$\max_{|z|\leq 1} |p_n(z; \lambda_\ell, \epsilon)| \leq \exp\left\{\frac{\theta}{n-\theta} + \frac{1}{2n} + \frac{1}{2}e{-}2\theta\right\} \leq \mathrm{e}^{1/d}, \tag{55}$$

where the concluding inequality is readily given by the choice of θ and by (54). Recall that by (44), $p_n(z; \lambda_\ell, \epsilon)$ is divisible by $(1 - \lambda_\ell\zeta)$; when setting

$$r_\ell(\zeta) = \frac{p_n(z; \lambda_\ell, \epsilon)}{1 - \lambda_\ell\zeta} = \sum_{\ell=0}^{n-1}[(1-\epsilon)\lambda_\ell\zeta]^\ell[1 - \ell/n],$$

and $r(\zeta) = \prod_{\ell=1}^d r_\ell(\zeta)$, we clearly have $r(0) = 1$, $\deg r \leq d(n-1) \leq m - d$, and

$$1 - q(\zeta) = p_{\mathbf{u}}(\zeta)\prod_{\ell=1}^d r_\ell(\zeta) = p_{\mathbf{u}}(\zeta)r(\zeta),$$

as required in (18) (note that q is a real polynomial due to $\mathbf{u} \in \Omega_d$).

2.2^0. By (55) we have $\|p_n(\cdot; \epsilon, \lambda_k)\|_\infty \leq \mathrm{e}^{1/d}$, while (53) says that

$$\|p_n(\cdot; \lambda_k, \epsilon) - 1\|_2 \leq 3\sqrt{\ln(2d)/n}.$$

We have

$$1 - \prod_{k=1}^{\ell+1} p_n(\cdot; \lambda_k, \epsilon) = \left[1 - \prod_{k=1}^{\ell} p_n(\cdot; \lambda_k, \epsilon)\right] p_n(\cdot; \lambda_{\ell+1}, \epsilon) + [1 - p_n(\cdot; \lambda_{\ell+1}, \epsilon)].$$

When denoting $\alpha_\ell = \|1 - \prod\limits_{k=1}^{\ell} p_n(\cdot; \lambda_k, \epsilon)\|_2$ for $\ell = 1, 2, ..., d$ and setting $\alpha_0 = 0$, we get

$$\alpha_{\ell+1} \leq \alpha_\ell\|p_n(\cdot; \lambda_{\ell+1}, \epsilon)\|_\infty + \|p_n(\cdot; \lambda_{\ell+1}, \epsilon) - 1\|_2 \leq \alpha_\ell\mathrm{e}^{1/d} + 3\sqrt{\ln(2d)/n},\ 0 \leq \ell < d.$$

It follows that $\alpha_\ell \leq 3\ell\sqrt{\frac{\ln(2d)}{n}}\mathrm{e}^{\ell/d}$ when $\ell \leq d$, which implies

$$\|q(\cdot)\|_2 \leq 3ed\sqrt{\ln(2d)/n} \leq 3\mathrm{e}d^{3/2}\sqrt{\frac{\ln(2d)}{m}},$$

and (19) follows. $\square$